# Parameter identification using the level set method

Zhiming Lu<sup>1</sup> and Bruce A. Robinson<sup>1</sup>

Received 22 December 2005; revised 6 February 2006; accepted 9 February 2006; published 22 March 2006.

[1] This study describes an inverse approach for efficiently identifying the spatial shapes of zones of low (or high) permeability using the level set method, given a set of spatially distributed head measurements. By this method, the boundaries of zones are characterized by a level set function. From an initial setting, the unknown regions of zones are determined by evolving the boundaries in artificial time using a pseudo velocity field that is related to the sensitivity of head to permeability and the residual between the measured head and modeled head at the current time. A synthetic example presented to illustrate the method. **Citation:** Lu, Z., and B. A. Robinson (2006), Parameter identification using the level set method, *Geophys. Res. Lett.*, 33, L06404, doi:10.1029/2005GL025541.

#### 1. Introduction

[2] Identifying parameter zonations is probably the most difficult in parameter identification problems. Traditionally, the heterogeneous domain of interest is divided into a number of zones and the parameter value in each zone is a constant to be determined. Although boundaries of these zones have significant impact on predicting flow and solute transport in the domain, in most cases we do not have enough direct information to infer the size, shape, locations, and the number of zones. Even in cases for which there is a clear correlation between identifiable geologic indicators and hydraulic conductivity, often data control is still insufficient to infer the size, shape, and location of zones. More problematic is the situation in which hydraulic conductivity does not correlate well with lithology. The zonation problem is extremely ill-posed in these cases. Sun and Yeh [1985] were the first to propose a method to identify simultaneously both the parameter zonation and its parameter values for the hydraulic conductivity field. Using some model structure identification criteria, Carrera and Neuman [1986] were able to choose the best parameter zonation pattern among a number of given alternatives. Eppstein and Dougherty [1996] used a modified version of the extended Kalman filter, a data-driven procedure that dynamically determines and refines zonations. Tsai et al. [2003] used Voronoi zonation to parameterize the unknown distributed parameter and solved the inverse problem by a sequential globallocal optimization procedure.

[3] In this study, we introduce a new approach for parameter zonation identification based on the level set method, applying the approach to a simple case of one material embedded in another. This method can be used to

This paper is not subject to U.S. copyright. Published in 2006 by the American Geophysical Union. identify, for example, low-permeability layers in a relatively higher permeability porous media (or vice versa), or highly permeable fault zones in the subsurface.

[4] The level set method is a very powerful tool for solving problems that involve geometry and geometric evolution [Osher and Sethian, 1988]. It has also been applied to solving shape optimization problems [Burger, 2003]. By a shape we mean a bounded region  $D \in \mathbb{R}^n$  with a  $C^1$  boundary. Instead of working on D directly, in the level set method a function  $\phi(\mathbf{X})$ , with  $D = {\mathbf{X}, \phi(\mathbf{X}) < 0}$ , is manipulated to adjust D implicitly. Since D is unknown, so too is the function  $\phi(\mathbf{X})$ . In shape optimization problems we start from an initial shape and improve it iteratively, by updating an initial level set function  $\phi(\mathbf{X})$  iteratively. The method has been used in several fields, including image segmentation [Lie et al., 2005] and inverse problems [Santosa, 1996]. One of the advantages of the level set method is that it is much easier to work with a globally defined function than to keep track of the boundaries of regions of interest, which may split into many regions or merge into larger ones.

[5] It is important to emphasize that, comparing with geostatistical inverse methods such as indicator (co-)kriging, the inverse approach based on the level set method requires no *a priori* assumptions on shape, size and locations of zones to be sought or correlation structures of these zones. This advantage should be very useful for ill-posed problems in hydrogeology.

# 2. Problem Statement

[6] Consider transient water flow in saturated media satisfying the standard governing equation

$$\nabla \cdot [K_s(\mathbf{X})\nabla h(\mathbf{X},t)] + g(\mathbf{X},t) = S_s \partial h(\mathbf{X},t) / \partial t, \quad \mathbf{X} \in \Omega$$
 (1)

subject to appropriate initial and boundary conditions. Here  $h(\mathbf{X}, t)$  is the hydraulic head,  $K_s(\mathbf{X})$  is the saturated hydraulic conductivity,  $S_s$  is the specific storage, and  $\Omega$  is the flow domain of interest. For simplicity,  $S_s$  is taken to be constant, because its variation is relatively small compared to that of the hydraulic conductivity.

[7] To introduce this method in the simplest way possible, we assume that the saturated hydraulic conductivity is a spatially varying binary random variable, i.e., one material being (disjointly) embedded in the other. Although there is no direct information regarding the size, shape, and locations of these zones, it is assumed that the hydraulic conductivity values for these two materials are known. This assumption may be justified. In fact, in many sites, hydraulic conductivity values for different stratigraphic units are known or can be estimated, but the exact spatial distribution of units are not.

**L06404** 1 of 4

<sup>&</sup>lt;sup>1</sup>Hydrology, Geochemistry, and Geology Group (EES-6), Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

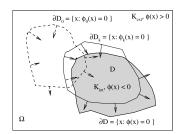


Figure 1. Schematic diagram showing the evolution of level set functions.

[8] Now suppose head values are measured at  $n_h$  locations  $\mathbf{X}_1$ ,  $\mathbf{X}_2$ ,  $\cdots$ ,  $\mathbf{X}_{n_h}$ , and at  $n_t$  times  $t_1$ ,  $t_2$ ,  $\cdots$ ,  $t_{n_t}$ , arranged in a vector  $h_0 = (h_i^{(k)})$ ,  $i = \overline{1, n_h}$ , and  $k = \overline{1, n_t}$ . The aim is to find the spatial distribution of two materials in the domain. To demonstrate the method, the case of error-free head measurements is examined, though it would be straightforward to include measurement errors in future work.

# 3. Level Set Representation of Zones

[9] Suppose that there is an unknown subset of  $\Omega$ , denoted as D, representing the regions of low- (or high-) permeability zones. The boundary of D can be described by a function  $\phi(\mathbf{X})$ ,  $\partial D = \{\mathbf{X}: \phi(\mathbf{X}) = 0\}$ . Accordingly, the hydraulic conductivity field can be represented by a binary function

$$K(\mathbf{X}) = \begin{cases} K_{int}, & \text{for } \mathbf{X} \in D = \{\mathbf{X} : \phi(\mathbf{X}) < 0\}, \\ K_{ext}, & \text{for } \mathbf{X} \in D^C = \{\mathbf{X} : \phi(\mathbf{X}) > 0\}, \end{cases}$$
(2)

where  $D^C = \Omega \backslash D$  is the complement of D, and  $K_{int}$  and  $K_{ext}$  are the hydraulic conductivities of interior and exterior of D, respectively. Note that for a given D there are an infinite number of functions that satisfy  $\partial D = \{\mathbf{X}: \ \phi(\mathbf{X}) = 0\}$ . However, any function  $\phi(\mathbf{X})$  uniquely defines  $\partial D$ . This feature allows us to reinitiate  $\phi(\mathbf{X})$  periodically without affect the boundary  $\partial D$ .

- [10] The inverse problem described above is to find  $\phi(\mathbf{X})$  such that the hydraulic head field solved using the spatial distribution of different zones characterized by  $\partial D = \{\mathbf{X}: \phi(\mathbf{X}) = 0\}$  matches head measurements at the various observation points and times. Note that no assumption has been made on the connectedness of D, i.e., D could be a disjoint set, or connected but including a number of holes. In addition, there is no assumption made regarding the number of zones, their size and locations, correlation structure, or the proportion of two materials.
- [11] Because the shape, size, and locations of set D are unknown, function  $\phi(\mathbf{X})$  is also unknown. In the level set approach, we generate a sequence of functions  $\phi_k(\mathbf{X})$ , defining a sequence of regions  $D_k = \{\mathbf{X}: \phi_k(\mathbf{X}) < 0\}$ , such that  $D_k \to D$ , as illustrated in Figure 1. In this iterative process, the number of zones and their shape, size, and locations are sequentially improved. The success of this method hinges on finding a strategy for efficiently propa-

gating the boundary of  $D_k$  such that it approaches D. This is described in the following sections.

#### 4. Formulation of the Inverse Problem

[12] The inverse problem described above can be written as a problem of minimizing an objective function:

$$F(K) = 0.5||H(K) - h_0||^2$$
(3)

where  $h_0$  is a vector of head measurements, and H represents an operator that maps the hydraulic conductivity field K to the hydraulic head field (i.e., the model solving the flow equation with appropriate boundary and initial conditions) and takes head samples at observation points and times.

[13] The directional derivative of function F(K) in the direction  $\delta K$ , the variation of K, is given by

$$\delta F(K) = \langle J^{T}(K)[H(K) - h_0], \delta K \rangle, \tag{4}$$

where J(K) is the Jacobian of the head to hydraulic conductivity, and  $\langle \cdot \rangle$  represents the inner product:

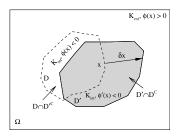
$$\delta F(K) = \int_{D} J^{T}(K)[H(K) - h_{0}] \delta K d\mathbf{X}.$$
 (5)

In the discretized case, J is an  $(n_h \times n_t) \times N$  Jacobian matrix, whose components are  $J_{ij} = dh_i/dK_j$ ,  $i = \overline{1, n_h \times n_t}$ ,  $j = \overline{1, N}$ , where N is the number of grid nodes in the domain  $\Omega$ .

[14] For a given variation  $\delta \phi$  (or  $\delta \mathbf{X}$ ) that propagates D to a new set D', as shown in Figure 2,  $\delta K$  is non-zero only in  $D'' = [D \cap D'^C] \cup [D' \cap D^C]$ , which represents those points that are either in D but not in D' or in D' but not in D. As a result, the integral in (5) can be reduced to an integral over D''. For an infinitesimal  $\delta \mathbf{X}$ ,  $D'' = \partial D$ , and (5) can be written as

$$\delta F(K) = \int_{\partial D} J^{T}(K) [H(K) - h_0] \delta K d\mathbf{X}.$$
 (6)

Certainly, variation  $\delta K$  is caused by variation  $\delta \mathbf{X}$  (or  $\delta \varphi$ ), and we need to represent  $\delta K$  in terms of  $\delta \mathbf{X}$  or  $\delta \varphi$ . Following *Santosa* [1996], as illustrated in Figure 2,  $\delta K$  can be related to  $\delta \mathbf{X}$  by  $\delta K = (K_{int} - K_{ext})\mathbf{n}(\mathbf{X}) \cdot \delta \mathbf{X}|_{\mathbf{X} \in \partial \mathbf{D}}$ , where  $\mathbf{n}(\mathbf{X}) = \nabla \varphi(\mathbf{X})/|\nabla \varphi(\mathbf{X})|$  is the normal direction of the curve at  $\mathbf{X}$ . We can assume that each



**Figure 2.** Schematic diagram illustrating the relationship between  $\delta \mathbf{X}$  and  $\delta K$ , which is non-zero only in  $D \cap D'^C$  or  $D' \cap D^C$ .

point on  $\partial D$  moves in the normal direction of the boundary  $\partial D$ ,  $\delta \mathbf{X} = \alpha(\mathbf{X})\mathbf{n}(\mathbf{X})$ , where  $\alpha(\mathbf{X})$  can be viewed as the propagation speed of the boundary. Substituting expressions of  $\delta \mathbf{X}$  and  $\delta K$  into (6),

$$\delta F(K) = \int_{\partial D} J^{T}(K) [H(K) - h_0] (K_{int} - K_{ext}) \alpha(\mathbf{X}) d\mathbf{X}$$
 (7)

To ensure that  $\delta F(K)$  is negative,  $\alpha$  is chosen as

$$\alpha(\mathbf{X}) = -sgn(K_{int} - K_{ext})J^{T}(K)[H(K) - h_{0}], \tag{8}$$

where sgn is the sign function, which returns the sign of its argument. Since the boundaries are propagated at speed  $\alpha$ , the direct measurements of the saturated hydraulic conductivity can be incorporated easily by setting the initial field to honor the measured hydraulic conductivity and then forcing the propagation speed to be zero at these measurement locations such that the boundaries will not across these locations.

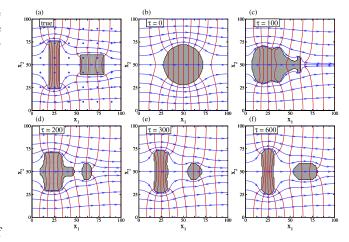
[15] The physical meaning of (8) can be explained as follows. With no loss of generality, consider the case in which  $K_{int} < K_{ext}$ . Suppose that at point **X** at a given iteration, the modeled head values are too high compared to the measured values (i.e.,  $H(K) > h_0$ ), and that an increase of hydraulic conductivity value at **X** will increase the head at the observation points (i.e., J(K) > 0). In this case, the  $\alpha$  value at point **X** computed from (8) will be positive, signifying that the boundary  $\partial D$  at point **X** should move outward. As a result, the hydraulic conductivity value at **X** decreases from  $K_{ext}$  to  $K_{int}$ , which will lead to a decrease in modeled head values in the new iteration.

[16] Note that the propagation speed  $\alpha(\mathbf{X})$  combines contributions from all observation points. If the sensitivity values of head at observation points  $\mathbf{X}_i$  and  $\mathbf{X}_j$  to hydraulic conductivity at  $\mathbf{X}$  are the same, the point with larger head residual will exert more influence on the propagation speed in the next iteration. On the other hand, if the head residuals at two observation points (or times) are the same, the point with a higher sensitivity value will be more influential. Finally, should the modeled head match the observed head, the boundaries would stop evolving.

[17] The final step is to derive an equation for the evolution of  $\phi$ . The variation of  $\mathbf{X}$  can be related to that of  $\phi$  by taking the variation of equation  $\phi(\mathbf{X}) = 0$ ,  $\delta\phi(\mathbf{X}) + \nabla\phi(\mathbf{X}) \cdot \delta\mathbf{X} = 0$ . If the function  $\phi(\mathbf{X})$  is expressed as a function of both  $\mathbf{X}$  and an artificial time  $\tau$ ,  $\phi = \phi(\mathbf{X}, \tau)$ , the evolution of  $\phi(\mathbf{X}, \tau)$  accordingly defines the evolution of boundary  $\partial D(\tau) = \{\mathbf{X}: \phi(\mathbf{X}, \tau) = 0\}$ . For sufficiently large  $\tau$ ,  $\partial D = \{\mathbf{X}: \phi(\mathbf{X}, \tau) = 0\}$  defines the spatial distribution of two materials and gives the solution to the inverse problem. Substituting expressions of  $\delta\mathbf{X}$  into above variational relationship and recalling the definition of  $\mathbf{n}(\mathbf{X})$ , an initial value problem for  $\phi(\mathbf{X}, \tau)$  is derived,

$$\partial \phi(\mathbf{X}, \tau) / \partial \tau + \alpha(\mathbf{X}, \tau) |\nabla \phi(\mathbf{X}, \tau)| = 0, \phi(\mathbf{X}, 0) = \phi_0(\mathbf{X}), \quad (9)$$

which is called the level set equation. The solution of this equation over the artificial time  $\tau$  gives the boundary  $\partial D(\tau)$ 



**Figure 3.** Comparison of the true distribution of low-permeability zones with inverse results at different stages of computation.

and eventually the solution stabilizes, representing the solution to the inverse problem.

# 5. Calculation of the Jacobian Using the Adjoint Method

[18] The derivative of hydraulic head at each measurement location with respect to hydraulic conductivity at each grid node in the domain  $\Omega$  can be written as

$$dh(\mathbf{X}_{i},t)/dK_{j} = \int_{0}^{t} \int_{\Omega_{j}} \nabla H(\mathbf{X},t) \cdot \nabla \left[ \partial \psi(\mathbf{X},t) / \partial t |_{t-\tau} \right] d\mathbf{X} d\tau,$$
(10)

where H is the mean head  $K_j$  is the saturated hydraulic conductivity at grid node j,  $\Omega_j$  is the exclusive subdomain of node j, and  $\psi$  is the adjoint state variable, which can be solved from the following adjoint state equation:

$$\nabla [K_G(\mathbf{X})\nabla \psi(\mathbf{X},t)] + \delta(\mathbf{X} - \mathbf{x}_i) = S_s \partial \psi(\mathbf{X},t) / \partial t, \tag{11}$$

with homogeneous initial and boundary conditions. Here  $X_i$  is an observation point and  $\delta()$  represents the Dirac  $\delta$  function. Equation (11) needs to be solved  $n_h$  times by placing an injection well of unit strength at each observation point.

# 6. Illustrative Examples

[19] In this section, we illustrate the level set method with a two-dimensional synthetic example. The test system consists of steady-state, saturated water flow in a rectangular domain of  $100 \text{ m} \times 100 \text{ m}$ , discretized into elements of a size  $1 \text{ m} \times 1 \text{ m}$ . The true (synthetic) permeability field, shown in Figure 3a, consists of two zones of low-permeability material ( $k = 10^{-13} \text{ m}^2$ ), one tall vertical bar and one fat horizontal bar, embedded in a background material of  $k = 10^{-10} \text{ m}^2$ . The boundary conditions are prescribed as constant head at left ( $H_1 = 10.5 \text{ m}$ ) and right ( $H_2 = 10.0 \text{ m}$ ) boundaries and no flow at the two lateral boundaries. The steady-state flow equation is solved for the synthetic permeability field, and 36 head measurements at

various locations are assumed to be available, shown as points in Figure 3a. Figure 3a also shows the contour lines and streamlines for the true flow field.

[20] We then compute the Jacobian matrix, which is the derivative of the hydraulic head at observation points  $\mathbf{X}_i$  with respect to permeability at each node at the domain  $\Omega$ , using (10)–(11). Note that the Jacobian is a function of the mean hydraulic head  $H(\mathbf{X},t)$ , which in turn depends on the mean permeability field. Ideally, the Jacobian matrix should be updated at each iteration. Due to high computational cost, however, it is calculated only once using the mean head solving from a homogeneous permeability field of  $k=10^{-10}$  m<sup>2</sup>.

[21] The iterative procedure is initiated by choosing an initial level set function  $\phi_0(\mathbf{X}) = 0.1 - \exp(-0.005d)$ , where  $d = (x - 50)^2 + (y - 50)^2$  is the distance from point  $\mathbf{X} = (x, y)$  to the center of the domain (50 m, 50 m). This function defines an initial guess of low-permeability zones (a disc with radius of 21.46 m, located at the center of the domain, see Figure 3b). The flow equation is then solved using the initial permeability field defined by  $\phi_0(\mathbf{X})$ , and the pseudo velocity field  $\alpha(X)$  is computed based on the difference between the modeled head and the observed head using (8). Finally, the boundary is propagated by solving the level set equation (9). The new boundary of low-permeability zones is determined by the set  $\partial D =$  $\{X: \phi(X) = 0\}$ . This process is repeated until either the prescribed number of updates has been reached or the difference between the modeled head and the observed head is smaller than a prescribed value. To prevent steep or flat gradients of the level set function  $\phi$ , the function is periodically reinitialized to a signed distance function, i.e.,  $\phi(\mathbf{X}) = -d(\partial D, \mathbf{X})$  for  $\mathbf{X} \in D$  or  $\phi(\mathbf{X}) = d(\partial D, \mathbf{X})$ for  $X \in \Omega \backslash D$ , where d represents the distance between point **X** and boundary  $\partial D$ . Note that this reinitialization will not change the boundary  $\partial D$ .

[22] The evolution of the boundary  $\partial D$  is depicted in Figures 3c-3f, where the artificial time  $\tau$  represents the number of updates. The method obtains the inversion results that are very similar to the true field. The algorithm proceeds to form the vertical bar first because it has a more dominant impact on the flow field. After resolving this bar, the algorithm then forms the fat horizontal bar. To ensure the convergence of the algorithm, a relatively small time step has been used, and the model run takes about 600 steps. The number of required steps may be reduced if the length of the time step varies dynamically. This demonstration shows that the level set method can efficiently handle the splitting or merging of the regions, provided the underlying data support is adequate for developing a model with this resolution.

[23] To investigate the sensitivity of the inverse model to the initial setting, we conducted two more model runs, one with the same size of inclusion in the upper-left corner of the domain rather than in the center as in the previous example and the other run with an initial inclusion of diameter 10 units in the upper-right corner of the domain. The inversion results from the first run are almost identical to the results illustrated in Figure 3f, while the second run does not capture any inclusions in the domain. A general

rule is that the initial setting should be chosen such that it has a significant impact on the flow field.

# 7. Summary and Further Work

[24] This paper is a preliminary study proposing a new inverse algorithm based on the level set method. In this method, the boundary of low- (or high-) permeability zones is represented by the zero level function. Starting from an initial setting of zones, their boundary is sequentially evolved to reduce the difference between the observed hydraulic head and the modeled head. The propagation speed of the boundary at any iteration is proportional to the sensitivity of head to the permeability field and the residual between the observed and modeled head at various measurement locations and observation times. The synthetic example presented showed that the level set method can efficiently identify the parameter zonation.

[25] These promising initial results suggest that further work is warranted to explore the level set method in more detail. Future work will include extending the method to incorporate transient head response data; using other data sets beyond simply head data (solute concentration or travel time measurements); consideration of measurement error and data density into the evaluation of the method; and developing methods for the joint inversion for shape and permeability of the individual zones. Furthermore, for maximum usefulness, the method should be extended to account for an arbitrary number of stratigraphic units, each with a distinct unknown value of permeability. Finally, one important issue that needs to be addressed in the future is the uncertainty associated with the prediction of internal boundaries between different materials, which may provide direct input to the random domain decomposition method of Winter and Tartakovsky [2000] for better quantifying flow and solute transport in the subsurface.

# References

Burger, M. (2003), A framework for the construction of level set methods for shape optimization and reconstruction, *Interface Free Boundaries*, 5, 301–329.

Carrera, J., and S. P. Neuman (1986), Estimation of aquifer parameters under transient and steady sate conditions: 3. Application to synthetic and field data, *Water Resour. Res.*, 22(2), 228–242.

Eppstein, M. J., and D. E. Dougherty (1996), Simultaneous estimation of transmissivity values and zonation, *Water Resour. Res.*, 32(11), 3321–3336

Lie, J., M. Lysaker, and X. Tai (2005), A piecewise constant level set framework, *Int. J. Numer. Anal. Model.*, 2, 422–438.

Osher, S., and J. Sethian (1988), Front propagation with curvature dependent speed: Algorithm based on Hamilton-Jacobi formulations, J. Comput. Phys., 56, 12–49.

Santosa, F. (1996), A level set approach for inverse problems involving obstacles, *ESAIM Control Optim. Calculus Variations*, *1*, 17–33.

Sun, N.-Z., and W. W.-G. Yeh (1985), Identification of parameter structure in groundwater inverse problem, *Water Resour. Res.*, 21(6), 869–883.

Tsai, F. T.-C., N.-Z. Sun, and W. W.-G. Yeh (2003), Global-local optimization for parameter structure identification in three-dimensional groundwater modeling, *Water Resour. Res.*, 39(2), 1043, doi:10.1029/2001WR001135.

Winter, C. L., and D. M. Tartakovsky (2000), Mean flow in composite porous media, *Geophys. Res. Lett.*, 27(12), 1759–1762.

B. A. Robinson and Z. Lu, Hydrology, Geochemistry, and Geology Group (EES-6), MS T003, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. (zhiming@lanl.gov)